

Use and Evaluation of an Expert System for Interpretation of Oral Radiographic Lesions.

W. Paul Lang, D.D.S., M.P.H. and Sharon L. Brooks, D.D.S., M.S.
University of Michigan School of Dentistry
Ann Arbor, Michigan 48109-1078.

INTRODUCTION

Use and evaluation of decision-support systems should be integral components of dental education. A senior level dental informatics course required students to generate differential diagnoses with an expert system, Oral Radiographic Differential Diagnosis (ORAD) [1] and to evaluate the system. In ORAD, users answer 16 questions regarding radiographic characteristics of jaw lesions such as location, size, borders, contents, and effect on adjacent teeth. A Bayesian statistical approach then analyzes 140 lesions resulting in a differential diagnosis and a likely probability for the observed lesion.

METHODS

Students (N = 89) were randomly assigned two radiographic jaw lesions to diagnose following a brief in-class demonstration of ORAD. Exercise instructions and 24 case studies with radiographs were made accessible to students via a WWW server. ORAD software was loaded independently on Macintosh computers in a computing laboratory. Students accessed the cases, opened ORAD, then used a customized Web form to enter their diagnoses and evaluative comments. At the completion of the exercise, the form was printed and submitted for grading.

Students' diagnoses were evaluated by an oral radiologist (SLB) who developed two likely diagnoses for each lesion using ORAD and determined a 'gold standard' diagnosis using either her expert judgment or biopsy results. Since ORAD could generate an extensive list of diagnoses with diminishing probabilities of importance, the students' top two diagnoses were used to evaluate their skills with ORAD.

RESULTS

A total of 85 cases (96%) was analyzed. On average, each lesion was assessed by students 7.0 times (Range 5-8). The mean number of different diagnoses generated per case across all students was 26.0 (Range = 13-36). The average number of first or second diagnoses per case was 7.1 (Range 3-11). The overall percentage of first or second student diagnoses considered to be acceptable by the radiologist was 49.0% (Range = 0.0-100.%).

Using ORAD on all cases, the radiologist developed 70.8% correct diagnoses, 12.5% partially correct diagnoses, and 16.7% incorrect diagnoses.

Student evaluations of ORAD were generally favorable. Positive comments included: user friendly software, responds quickly, self-explanatory, fits current therapeutic and diagnostic problem solving, explores possibilities the operator may have overlooked, and clear explanatory facility. Negative comments included: system is oversimplified and difficult to manipulate, changing one criterion radically changes the differential diagnoses, limited choices of signs and symptoms are available, a book would be easier to use, and key questions are missing from the question option list.

About 20% of the students' cases were not correctly submitted on the customized Web form. On a scale of 1 = Strongly Agree to 5 = Strongly Disagree, the average class response to the question "The ORAD exercise was a useful part of the informatics course." was 2.0.

DISCUSSION

The diagnosis lists developed by students were very broad, and their success with diagnosing lesions was lower than expected. Reasons for these outcomes could be students' unfamiliarity with ORAD, absence of some lesions from ORAD's database, incorrect interpretation of lesion features, and some poor quality Web images. Future plans include comparisons of students' diagnostic success using ORAD and standard methods of diagnosis, pre-assignment ORAD tutorials, and improvement of Web images. Possible ORAD activities include limiting its generation of extensive diagnostic lists, enlarging its database, and making diagnostic comparisons of ORAD and expert radiologists with gold standard lesions. Also, use and explanation of Web forms require further development.

CONCLUSIONS

Students' diagnostic success with ORAD was low, although their attitudes toward the expert system exercise were favorable. ORAD produced somewhat reasonable diagnoses when used by an expert radiologist. Expansion of ORAD's lesion database could increase the likelihood of correct diagnoses. Improved development and presentation of Web-based cases and electronic form submission are needed.

REFERENCES

- [1]. White SC. Computer-aided differential diagnosis of oral radiographic lesions. *Dentomaxillofac Radiol* 1989;18:53-9.